P3.4

MULTIVARIATE FORECAST ERROR COVARIANCES FOR AN OCEAN MODEL ESTIMATED BY MONTE-CARLO SIMULATION.

Anna Borovikov*
Department of Mathematics
University of Maryland, College Park, Maryland

Michele M. Rienecker
Oceans and Ice Branch, Laboratory for Hydrospheric Processes
NASA/Goddard Space Flight Center, Greenbelt, Maryland

1. INTRODUCTION

One of the most difficult aspects of ocean state estimation is the prescription of the model forecast error covariances. The paucity of ocean observations limits our ability to estimate the covariance structures from model-data differences. In most practical applications, simple covariances are usually prescribed. Rarely are cross-covariances between different model variables used, and, for example, only the temperature field is analyzed in a univariate Optimal Interpolation (OI). However, it has been found that a univariate OI has a detrimental effect on the salinity and velocity fields of the model (e.g., Troccoli et al, 2001). Apparently, in a sequential framework it is important to analyze temperature and salinity together. Here an estimation of the model error statistics is made by Monte-Carlo techniques from an ensemble of model integrations. An important advantage of using an ensemble of ocean states is that it provides a natural way to estimate cross-covariances between the fields of different physical variables constituting the model state vector. This study gives the details of the two assimilation experiments different only in the model error covariance specification, thus allowing for comparison of the OI performance with a traditional Gaussian model of the error covariance and an empirical multi-variate model. Robustness of the empirical multi-variate error covariance estimate is explored.

2. OI ASSIMILATION

In this study the OI sequential data assimilation scheme was used for the ocean state estimation. A distinct feature of the OI algorithm is that it uses fixed forecast error covariance estimates.

Corresponding author address: Anna Borovikov, Code 971, NASA/Goddard, Greenbelt, MD 20771, e-mail: *ayb@mohawk.gsfc.nasa.gov*

2.1 *Model and Data*

The Poseidon reduced-gravity quasi-isopycnal ocean model, introduced by Schopf and Loughe (1995), provides realistic simulations of tropical Pacific climatology and variability, as demonstrated, for example, in Borovikov et al (2001). For this study the domain was restricted to the Pacific Ocean (45°S to 65°N).

The TAO Array, consisting of more than 70 moored buoys spanning the equatorial Pacific (http://www.pmel.noaa.gov/toga-tao/home.html) measures subsurface temperature down to a depth of 500 meters. By 1994 daily measurements became available at approximately uniformly spaced locations across the equatorial Pacific Ocean. These are the sole source of data used in the assimilation experiments presented here. The standard deviation of observation measurement error is set to 0.5°C and the errors are assumed to be uncorrelated in space and time. This level of error is higher than the instrumental error in order to account for model representation errors.

For cross-validation, we use the meridional sections of temperature and salinity data from conductivity-temperature-depth (CTD) instruments processed by Johnson et al (2000).

3. FORECAST ERROR COVARIANCE MODELING

When attempting error covariance structure modeling, one is striving for both accuracy and efficiency. These two goals may often be counteracting. With limited knowledge of the true nature of the model error mainly due to the data sparsity, one often leans towards the efficiency criterion when choosing a particular implementation. We present here two different models for the forecast error covariance structure, a simpler and less computationally intensive and a more elaborate and hopefully more accurate model.

3.1 *Univariate Functional Error Covariance Model*

In the simple model, the spatial structure of the temperature (T) forecast error was assumed to be Gaussian in all three dimensions with scales 15°, 4° and 50 m in zonal, meridional and vertical directions correspondingly. During each assimilation cycle a correction was made only to the model temperature field. The other variables adjusted according to the model's response to the temperature correction. The assimilation experiment with this covariance model is denoted UOI.

3.2 *Monte Carlo Method for Estimating Multivariate Error Covariance*

In hopes of obtaining a more realistic covariance structure consistent with model dynamics, an application of the Monte Carlo method was proposed, that would use the variability from an ensemble of model integrations for a one-time estimate of the model forecast error statistics. In spirit, this approach is similar to the Ensemble Kalman Filter except that the error covariance does not evolve with time. Since the rank of the error covariance matrix P estimated using this method is no greater than the Monte Carlo ensemble size, it can be conveniently represented using a basis of empirical-orthogonal functions (eofs).

Of the approximately 150 eofs that were resolved as described below, the first 50 corresponding to the largest eigenvalues were retained for the forecast error covariance computation. The assimilation experiment with this forecast error model is denoted MvOI.

3.2.1. *Ensemble Generation*

An ensemble of ocean states was generated by forcing the ocean model with an ensemble of air-sea fluxes. These fluxes were obtained from a series of integrations of an atmospheric model forced by the same interannually varying sea surface temperatures (SST) and differing only in slight perturbations to the initial atmospheric state. The interannual anomalies in surface stress and heat flux components were added to climatological seasonal forcing derived from observations. This approach attributes all of the ocean model forecast error to uncertainties in the longer time scale surface flux anomalies, since differences between the ensemble members were due to atmospheric internal variability. In all, 32 ocean model runs were conducted. Five-day averages (pentads) of the model fields were saved. They were subsequently interpolated to 11 depth levels, coincident with the depths of the TAO observations. All the covariance estimates have been made using these fields.

4. ROBUSTNESS OF THE MULTIVARIATE ERROR COVARIANCE ESTIMATE

Initial experiments (not presented here) indicated that single snapshots from the 32-members of the ensemble yielded rather noisy covariance structures, particularly at large lags. We decided to artificially expand the ensemble size, by choosing snapshots from 5 years picked at random from each ensemble member. With appropriate means removed, the anomalies provided an ensemble 5 times the size of the number of individual simulations.

The procedure was repeated 10 times, and 10 sets of eigenvectors were compared. If a statistically significant difference were detected between these realizations of P, the forecast error covariance matrix, it would indicate the importance of interannual variations in the model error variability. To understand whether the various sets of eofs span the same dominant error subspace we considered the projection of an arbitrary collection of anomalous ocean states (prepared in the same way as for P) onto a given set of eofs. If the residual remaining after this projection is noise-like, the eofs captured the significant information regarding the model error covariance structure. An anomalous model ocean state vector a can be expressed (i.e. projected) in terms of the eof basis $\{\alpha\}$ as

$$a = \Sigma_i a_i \alpha_i + \delta^\alpha. \tag{1}$$

Here δ^α is the residual due to the imperfect representation of the space containing a by the space spanned by $\{\alpha\}$. In the preparation of the ensemble set of anomalous states a, care was taken not to include any sample used for $\{\alpha\}$. Thus, we obtain an ensemble of residuals $\{\delta\}$. We then compute the eofs of this residual ensemble to detect any modes of model error variability that were missing in the $\{\alpha\}$.

The results of 10 such tests are shown in figure 1: the eigenvalues corresponding to the original (model error covariance) eofs ($\{\alpha\}$) and the eigenvalues of the covariance matrix of the residuals ($\{\delta\}$) corresponding to an ensemble of residuals constructed for July. The eigencurves for January appear very similar and are not shown. The spectrum of the eigenvalues of $\{\delta\}$ is flat, which is characteristic of white noise. It appears that at least 30 eofs have significant information about the error structures and that the space spanned by these eofs is fairly invariant.

5. EXPERIMENTAL RESULTS

The multivariate error statistics were validated in the experiment that assimilated the TAO temperature data from July 1996 to December 1998. The experimental setup is as following. The model was spun-up for 10 years and then run with time-dependent forcing for 1988-1998 in all the experiments. The initial conditions and the forcing were identical in all the assimilation experiments. In addition to the data assimilation runs, a forced model integration without assimilation served as a base-line for assessing the assimilation performance.

This run is referred to as the control run. In every assimilation experiment, the daily subsurface temperature data was assimilated once a day. To alleviate the effects of the large shock on the model resulting from the intermittent assimilation, the incremental update technique was used, as described by Bloom et al (1996).

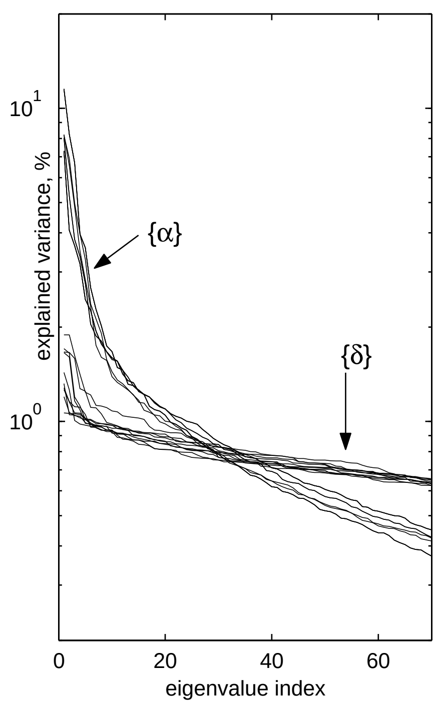

Figure 1: Eigenvalues for several realizations of the matrix P (marked $\{\alpha\}$) and the eigenvalues for ensembles of δ's - the residuals of the projections of an arbitrary collection of anomalous ocean states onto a basis of eofs. Only the first 75 of about 150 resolved eigenvalues are shown.

The longitude-time plot of the salinity field at 150 m (figure 2) shows a quick deterioration of the structure in UOI case. This is probably due to the fact that only the temperature field is corrected in UOI. The model dynamical balances are so much perturbed that the model is unable to restore them between assimilation cycles. The MvOI corrects all the model fields based on their empirical statistical relationship.

The assimilation is cross-validated by comparison of monthly-averaged T and S with CTD survey observations presented by Johnson et al (2000). These data have not been assimilated. The CTD sections spanned the equator from $8°S-8°N$. It is evident from the RMS errors between the model runs and CTD data shown in figure 3 that while both UOI and MvOI brought the temperature field closer to the observed, only the MvOI improves the salinity as well.

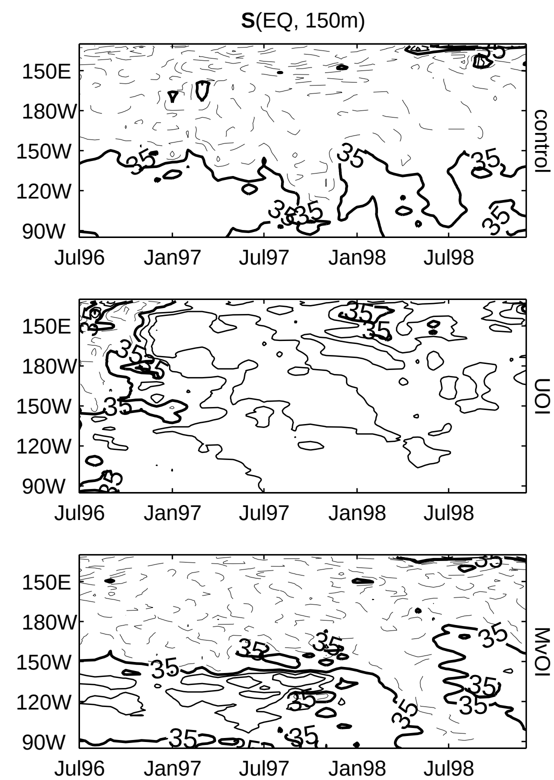

Figure 2: Time series of monthly model salinity field along the equator at 150 m. The average between $2°S$-$2°N$ is shown. The top panel shows the control run, the middle panel UOI, and the bottom panel, MvOI. Contours greater than 35 psu are dashed.

6. CONCLUSIONS

A multivariate OI algorithm, which allows for salinity and current fields to be corrected as a result of the temperature assimilation has been implemented. The multivariate algorithm uses anisotropic, inhomogeneous model error covariances obtained by a Monte Carlo simulation. In all, an ensemble of 160 members was constructed. The model error covariance matrix is naturally represented by a set of eofs, fewer in number than the ensemble size, allowing for an efficient analysis of its properties. The multivariate OI outperforms the univariate assimilation. The robustness of such an estimate was investigated and it was found that the model error covariance structure does not exhibit significant seasonal or interannual variations, although the latter should be taken cautiously in the view of a relatively short (15 years) analysis time interval. This result also indirectly validates the underlying assumption of the OI algorithm that the forecast error covariance structure is approximately constant.

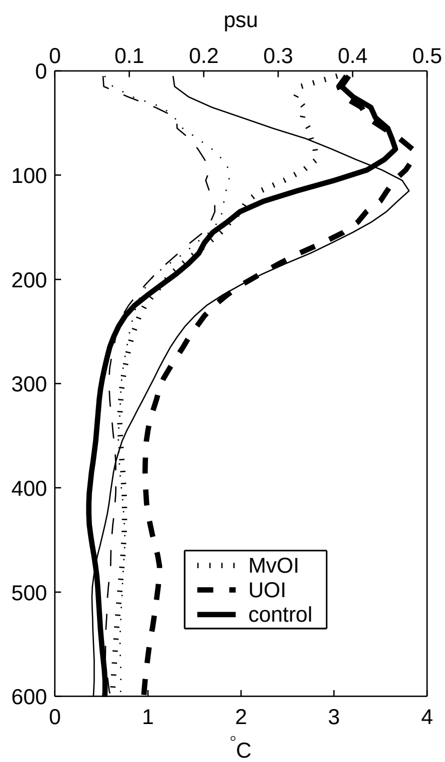

Figure 3: RMS difference between simulations and data for temperature and salinity, averaged over all available CTD profiles for the period 07/1996-12/1998 (Johnson et al., 2000). Thin lines correspond to temperature and thick lines to salinity. Salinity scale is at the top and temperature scale is at the bottom of the plot.

References

S. C. Bloom, L. L. Takacs, A. M. D. Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. Mon. Wea. Rev., **124**.

A. Borovikov, M. M. Rienecker, and P. S. Schopf, 2001: Surface heat balance in the equatorial Pacific ocean: climatology and the warming event of 1994-95. J. of Clim., **14(12)**, 2624–2641.

G. C. Johnson, M. J. McPhaden, G. D. Rowe, and K. E. McTaggart, 2000: Upper equatorial Pacific ocean current and salinity variability during the 1996-1998 El Niño-La Niña cycle. J. of Geophys. Res., **105**, 1037–1053.

P. S. Schopf and A. Loughe, 1995: A reduced-gravity isopycnal ocean model: hindcasts of El Niño. Mon. Weath. Rev., **3**, 2839–2863.

A. Troccoli, M. M. Rienecker, C. L. Keppenne, and G. C. Johnson, 2001: Temperature data assimilation with salinity corrections: validation in the tropical Pacific ocean, 1996-1998. J. Geophys. Res. submitted.